

Quantized Network Coding for Correlated Sources

Mahdy Nabaee, *Student Member, IEEE*, Fabrice Labeau, *Member, IEEE*

Abstract

Non-adaptive joint source network coding of correlated sources is discussed in this paper. By studying the information flow in the network, we propose quantized network coding as an alternative for packet forwarding. This technique has both network coding and distributed source coding advantages, simultaneously. Quantized network coding is a combination of random linear network coding in the (infinite) field of real numbers and quantization to cope with the limited capacity of links. With the aid of the results in the literature of compressed sensing, we discuss theoretical and practical feasibility of quantized network coding in lossless networks. We show that, due to the nature of the field it operates on, quantized network coding can provide good quality decoding at a sink node with the reception of a reduced number of packets. Specifically, we discuss the required conditions on local network coding coefficients, by using restricted isometry property and suggest a design, which yields in appropriate linear measurements. Finally, our simulation results show the achieved gain in terms of delivery delay, compared to conventional routing based packet forwarding.

Index Terms

Linear network coding, distributed source coding, compressed sensing, restricted isometry property, ℓ_1 -minimization.

I. INTRODUCTION

Flexible, low cost, and long lasting implementation of wireless sensor networks has made them an unavoidable alternative for conventional wired sensing structures in a wide variety of applications,

M. Nabaee and F. Labeau are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, E-mail: m.nabaee@ieee.org, fabrice.labeau@mcgill.ca

This work was supported by Hydro-Quebec, the Natural Sciences and Engineering Research Council of Canada and McGill University in the framework of the NSERC/Hydro-Quebec/McGill Industrial Research Chair in Interactive Information Infrastructure for the Power Grid.

including medicine, transportation, and military [1]. As a relatively new technology, the trends and challenges are more felt in the networking aspects of communication than in the classic physical layer era [2]. One of the introduced challenges is the gathering of sensed data at a central node of the network, where delivery delay, precision, and robustness to network changes are emerging issues.

As the conventional way of transmission in the networks, packet forwarding via routing is widely used in different implementations of sensor networks. While it achieves capacity rates in the case of lossless networks [3], packet forwarding requires an appropriate routing [4] protocol to be run. In the case of correlated sources, distributed source coding [5], [6] on top of packet forwarding is proved to be optimal, in terms of achieved conditional information rates [7]. However, packet forwarding can lead to difficulties because of its need for queuing, and its slow adaptation to network changes, caused by deploying new node(s) or link failure(s).

These issues and a lot more have motivated the invention of network coding [3], as an alternative for packet forwarding in sensor networks [8], [9]. Specifically, network coding sends a function of incoming packets to the intermediate nodes, as opposed to sending their original content. Furthermore, the usage of random linear functions, also known as random linear network coding, is proved to be sufficient in lossless networks [10], [11]. Moreover, theoretical analysis shows that when network coding is used for transmission, no queuing is required to achieve the optimal information rates [3]. Network coding in lossy networks can result in improved achieved rate regions, compared to packet forwarding [12], [13].

Similar to packet forwarding, network coding can be separately applied on top of distributed source coding for correlated sources [14], [15]. On the other hand, one has to perform joint source network decoding in order to achieve optimal performance limits, which may not be feasible because of its computational complexity [15]. Sub-optimal solutions have been proposed to tackle this practicality issue [16]–[18], by using low density codes and sum product algorithm [19] for decoding. Similar to the distributed source coding, which requires knowledge of appropriate marginal rates at each encoder node, these approaches need some knowledge of correlation model of sources, at the encoders' side. This knowledge of appropriate rates may be a luxury in some cases, especially when it is changing over time and needs to be updated frequently. Hence, it is essential to study the possibility of developing a *non-adaptive* joint source network coding for such cases. In this paper, we aim to develop a non-adaptive random linear network coding for efficient joint distributed source network coding of correlated sources in sensor networks.

Recently, the idea of using compressed sensing [20], [21] and sparse recovery concepts in sensor networks has drawn attention [22]–[25]. For instance, in [26], [27], theoretical discussion on sparse recovery

of graph constrained measurements with an interest in network monitoring application is presented. Joint source, channel and network coding was also proposed in [28], where random linear mixing was proposed for compression of temporally and spatially correlated sources. In [29], practical possibility of finite field network coding of highly correlated sources was investigated, with the aid of low density codes and belief propagation based decoding. Unfortunately, a solid theoretical investigation on the feasibility of adopting sparse recovery in random linear network coding has not been done previously.

In our earlier work [30], we proposed non-adaptive joint source network coding of exactly sparse sources, with the aid of the results in compressed sensing literature. In this paper, we extend our work to the general case of correlated sources and discuss theoretical and practical aspects of having robust distributed compression.

A detailed description of data gathering scenario with our notations is presented in section II. In section III, we introduce and formulate our proposed quantized network coding, which is followed by its theoretical feasibility discussion using restricted isometry property, in section IV. In section V, we present the decoding algorithm used to recover quantized network coded packets, and derive a performance bound on its recovery error. Our simulation setup and results are presented in section VI. Finally in section VII, we conclude the paper by discussion on our proposed method and the ongoing works on this topic.

II. PROBLEM DESCRIPTION AND NOTATION

In this paper, we consider a lossless model of networks, for which the links have limited capacities. Although it may not be a perfect model in practical cases where the links have mutual interference, it still reflects the effect of such imperfectnesses when calculating single input single output capacity of the links. A future work may study the case of noisy networks of links, by understanding the effect of interference between the links.

As shown in Fig. 1, we represent the network by a directed graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} and \mathcal{E} are the sets of nodes (vertices) and directed edges (links). Each node, v , is from the finite sorted set $\mathcal{V} = \{1, \dots, n\}$ and each edge, e , is from the finite sorted set $\mathcal{E} = \{1, \dots, |\mathcal{E}|\}$. Further, each edge (link) can maintain a lossless transmission from $tail(e)$ to $head(e)$, at a maximum finite rate of C_e bits per use. We define the sets of incoming and outgoing edges of node v , denoted by $In(v)$ and $Out(v)$, respectively, as follows:

$$In(v) = \{e : e \in \mathcal{E}, head(e) = v\}, \quad (1)$$

$$Out(v) = \{e : e \in \mathcal{E}, tail(e) = v\}. \quad (2)$$

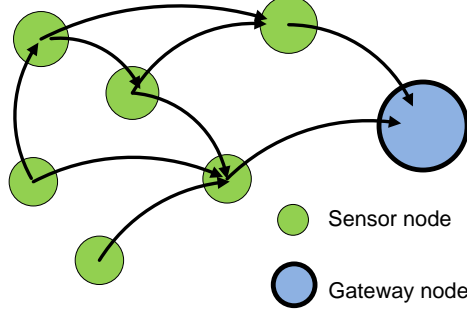


Fig. 1. Directed graph, representing a data gathering sensor network.

The input and output contents of edge e at time instant t are represented by $Y_e(t)$ and $Y'_e(t)$, where t represents the discrete (integer) time index, during which a block of length L is transmitted. Since the edges are lossless, $Y_e(t)$ and $Y'_e(t)$ are the same and from a finite alphabet of size $\lfloor 2^{LC_e} \rfloor$, where $\lfloor \cdot \rfloor$ denotes truncation to the lower integer. In the rest of the paper, the realizations of all capital letter random variables are denoted by lower case letters.

The nodes of the network are equipped with sensors and specifically each node v has an information source, X_v , where $X_v \in \mathbb{R}$. The sensed data are supposed to be correlated, as this is a valid assumption in a lot of different applications. We model the correlation between these sensed data, by the near-sparseness property, since it can be considered as a generalization of compressibility and sparseness. Specifically, by defining the sorted vector of X_v 's:

$$\underline{X} = [X_v : v \in \mathcal{V}], \quad (3)$$

we assume that \underline{X} is near-sparse in some orthonormal transform domain $\phi_{n \times n}$.¹ Explicitly, for $\underline{S} = \phi^T \cdot \underline{X}$, and a small positive ϵ_k , we have:

$$\|\underline{S} - \underline{S}_k\|_{\ell_1} \leq \epsilon_k, \quad (4)$$

where \underline{S}_k is such that:

$$\|\underline{S}_k\|_{\ell_0} = k, \quad (5)$$

and is called k -sparse. An example of the sparsifying transform matrix, ϕ , is the Karhunen Loeve transform of the messages.

Having these correlated information sources and the information network characterized, we study the

¹In this paper, all the vectors are column-wise.

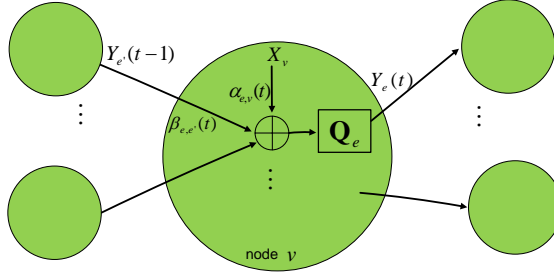


Fig. 2. A simple diagram of quantized network coding.

transmission of X_v 's to a single gateway node. The gateway or decoder node, denoted by v_0 , $v_0 \in \mathcal{V}$, has high computational resources and is usually in charge of forwarding the information to a next level network; *e.g.* a wired backbone network. The described (single session) incast of sources to the unique decoder node is referred to as *data gathering*. The purpose of this paper is to discuss the theoretical and practical feasibility of *non-adaptive* joint source network coding in the described data gathering scenario. More specifically, we take a compressed sensing approach in order to handle the transmission of sensed data.

III. QUANTIZED NETWORK CODING

Random linear network coding for multicast of independent sources has been proposed and studied in [11], where the algebraic operations are in finite field. Since our work is motivated by the concepts of compressed sensing, in which the results are valid in the infinite field of real number, we have to use a real field alternative for conventional finite field network coding. On the other hand, finite capacity of the edges has to be appropriately coped with the infinite number of symbols in the adopted real field network coding. As a result, we propose *Quantized Network Coding* (QNC), which uses quantization to bridge between the limited capacity of the links and infinite alphabet of real field network coded packets.

In [30], for $\forall v \in \mathcal{V}, \forall e \in \text{Out}(v)$, we defined QNC at node v , according to:

$$Y_e(t) = \mathbf{Q}_e \left(\sum_{e' \in \text{In}(v)} \beta_{e,e'}(t) \cdot Y_{e'}(t-1) + \alpha_{e,v}(t) \cdot X_v \right), \quad (6)$$

where $Y_e(1) = 0$, $\forall e \in \mathcal{E}$, ensures initial rest condition in the network. The messages, X_v 's are assumed to be constant until the transmission is complete.² The local network coding coefficients, $\beta_{e,e'}(t)$'s and $\alpha_{e,v}(t)$'s are real valued and are usually picked semi-randomly. The quantizer operator, $\mathbf{Q}_e(\cdot)$,

²This is why X_v 's do not depend on t .

corresponding to outgoing edge e , is designed based on the values of C_e and L , and the distribution of its input (*i.e.* random linear combinations). A simple diagram of QNC at node v is shown in Fig. 2.

Denoting the quantization noise of $\mathbf{Q}_e(\cdot)$ at time t , by $N_e(t)$, we can reformulate (6) as follows:

$$Y_e(t) = \sum_{e' \in In(v)} \beta_{e,e'}(t) \cdot Y_{e'}(t-1) + \alpha_{e,v}(t) \cdot X_v + N_e(t). \quad (7)$$

We denote the adjacency matrix, $[F(t)]_{|\mathcal{E}| \times |\mathcal{E}|}$, and $[A(t)]_{|\mathcal{E}| \times n}$ matrix, such that:

$$\{F(t)\}_{e,e'} = \begin{cases} \beta_{e,e'}(t) & , \text{tail}(e) = \text{head}(e') \\ 0 & , \text{otherwise} \end{cases}, \quad (8)$$

$$\{A(t)\}_{e,v} = \begin{cases} \alpha_{e,v}(t) & , \text{tail}(e) = v \\ 0 & , \text{otherwise} \end{cases}. \quad (9)$$

We also define the vectors of edge contents, $\underline{Y}(t)$, and quantization noises, $\underline{N}(t)$, according to:

$$\underline{Y}(t) = [Y_e(t) : e \in \mathcal{E}], \quad (10)$$

$$\underline{N}(t) = [N_e(t) : e \in \mathcal{E}]. \quad (11)$$

As a result, Eq. 7 can be re-written in the following form:

$$\underline{Y}(t) = F(t) \cdot \underline{Y}(t-1) + A(t) \cdot \underline{X} + \underline{N}(t). \quad (12)$$

Depending on the network deployment, matrix $[B]_{|In(v_0)| \times |\mathcal{E}|}$ defines the relation between the content of edges, $\underline{Y}(t)$, and the received packets at the decoder node. Explicitly, we define the vector of *marginal measurements* (received packets) at time t at the decoder:

$$\underline{Z}(t) = [Y_e(t) : e \in In(v_0)] = B \cdot \underline{Y}(t), \quad (13)$$

where:

$$\{B\}_{i,e} = \begin{cases} 1 & , i \text{ corresponds to } e, e \in In(v_0) \\ 0 & , \text{otherwise} \end{cases}. \quad (14)$$

By considering (12) as the difference equation, characterizing a linear system with \underline{X} and $\underline{N}(t)$'s as its inputs, and $\underline{Z}(t)$ its output, and using the results in [31], $\{Z(t)\}_i$'s are given by:

$$\underline{Z}(t) = \Psi(t) \cdot \underline{X} + \underline{N}_{\text{eff}}(t), \quad (15)$$

where the *marginal measurement* matrix, $\Psi(t)$, and the *marginal effective noise* vector, $\underline{N}_{\text{eff}}(t)$, are calculated as follows:

$$\Psi(t) = B \cdot \sum_{t'=2}^t \prod_{t''=t'}^{t'+1} F(t'') A(t'), \quad (16)$$

$$\underline{N}_{\text{eff}}(t) = B \cdot \sum_{t'=2}^t \prod_{t''=t'}^{t'+1} F(t'') \underline{N}(t'). \quad (17)$$

In Eqs. 16,17, the matrix multiplication is defined as:

$$\prod_{t''=t}^{t'+1} F(t'') = F(t) \cdots F(t' + 1). \quad (18)$$

By storing $\underline{Z}(t)$'s, at the decoder, we build up the *total measurements vector*, $\underline{Z}_{\text{tot}}(t)$, as follows:

$$\underline{Z}_{\text{tot}}(t) = \begin{bmatrix} \underline{Z}(2) \\ \vdots \\ \underline{Z}(t) \end{bmatrix}_{m \times 1}, \quad (19)$$

where $m = (t - 1)|\text{In}(v_0)|$. Therefore, the following can be established:

$$\underline{Z}_{\text{tot}}(t) = \Psi_{\text{tot}}(t) \cdot \underline{X} + \underline{N}_{\text{eff,tot}}(t), \quad (20)$$

where the $m \times n$ *total measurement matrix*, $\Psi_{\text{tot}}(t)$, and the *total effective noise* vector, $\underline{N}_{\text{eff,tot}}(t)$, are the concatenation result of marginal measurement matrices, $\Psi(t)$'s, and marginal effective noise vectors, $\underline{N}_{\text{eff}}(t)$. Because of our assumption to start transmission from $t = 1$, $\{\underline{Z}(1)\}_i$'s are not useful for decoding, and therefore:

$$\Psi_{\text{tot}}(t) = \begin{bmatrix} \Psi(2) \\ \vdots \\ \Psi(t) \end{bmatrix}, \quad (21)$$

$$\underline{N}_{\text{eff,tot}}(t) = \begin{bmatrix} \underline{N}_{\text{eff}}(2) \\ \vdots \\ \underline{N}_{\text{eff}}(t) \end{bmatrix}. \quad (22)$$

In the conventional linear network coding, the number of total measurements, m , is at least equal to the number of data, n . More precisely, the total measurement matrix is of full column rank, which makes

us able to uniquely find a solution.³ In this paper, we are interested to investigate the feasibility of robust recovery of \underline{X} , when fewer number of measurements are received at the decoder than the number of messages; *i.e.* $m < n$.

Considering the characteristic equation of (20), describing the QNC scenario, we can treat as a compressed sensing measurement equation. This gives us an opportunity to apply the results in the literature of compressed sensing and sparse recovery [20], [32] to our QNC scenario with near-sparse messages. However, one needs to examine the required conditions which guarantee sparse recovery in the proposed QNC scenario. In the following, we discuss theoretical and practical feasibility of robust recovery with a compressed sensing perspective.

IV. RESTRICTED ISOMETRY PROPERTY

One of the advantageous of compressed sensing is to tackle a non-adaptive design for sensing of sparse signals, where the support (location of non-zero elements) is not known at the encoding side. To pay back this non-adaptive characteristic, we may need more measurements than the exact number of non-zero elements. Fortunately, if appropriate types of linear measurements are chosen, we can keep the required number of measurements much less than the number of messages; that is: $m \ll n$.

One of the properties that is widely used to characterize appropriate measurement matrices in the compressed sensing literature, is the *Restricted Isometry Property* (RIP) [33]. Roughly speaking, it provides a measure of norm conservation while the dimensionality is reduced [34]. An $m \times n$ matrix $\Theta_{\text{tot}}(t)$ is said to satisfy RIP of order k with constant δ_k , if for all k -sparse vectors $\underline{s}_k \in \mathbb{R}^n$, we have:

$$1 - \delta_k \leq \frac{\|\Theta_{\text{tot}}(t)\underline{s}_k\|_{\ell_2}^2}{\|\underline{s}_k\|_{\ell_2}^2} \leq 1 + \delta_k. \quad (23)$$

Remark 4.1: Random matrices with identically and independently distributed (i.i.d) zero mean Gaussian entries are appropriate measurement matrices for compressed sensing. Explicitly, an $m \times n$ i.i.d Gaussian random matrix, denoted \mathbf{G} , with entries of variance $\frac{1}{m}$, satisfies RIP of order k and constant δ_k , with a probability exceeding

$$1 - e^{-\kappa_1 m}, \quad (24)$$

(called *overwhelming* probability) if

$$m > \kappa_2 k \log\left(\frac{n}{k}\right). \quad (25)$$

³beyond the fact that there should not be any uncertainty involved as a result of noise.

In (24), (25), κ_1 and κ_2 only depend on the value of δ_k (theorem 5.2 in [35]).

In [30], [36], we proposed a design for local network coding coefficients, $\beta_{e,e'}(t)$'s and $\alpha_{e,v}(t)$'s, which results in an appropriate total measurement matrix, $\Psi_{\text{tot}}(t)$, in the compressed sensing framework.

Theorem 4.2: (Theorem 3.1 in [36]) Consider a quantized network coding scenario, in which the network coding coefficients, $\alpha_{e,v}(t)$ and $\beta_{e,e'}(t)$, are such that:

- $\alpha_{e,v}(t) = 0, \forall t > 2$,
- $\alpha_{e,v}(2)$'s are independent zero mean Gaussian random variables,
- $\beta_{e,e'}(t)$'s are deterministic.

For such a scenario, the entries of the resulting $\Psi_{\text{tot}}(t)$ are zero mean Gaussian random variables. Further, the entries of different columns of $\Psi_{\text{tot}}(t)$, i.e. $\{\Psi_{\text{tot}}(t)\}_{i,v}$ and $\{\Psi_{\text{tot}}(t)\}_{i',v'}$, where $v \neq v'$, are independent. ■

It is also numerically shown in [36] that locally orthonormal set of $\beta_{e,e'}(t)$'s is a better choice than non-orthonormal sets; that is for all $e, e' \in \text{Out}(v)$, we have:

$$\sum_{e'' \in \text{In}(v)} \beta_{e,e''}(t) \cdot \beta_{e',e''}(t) = 0, \quad e \neq e', \quad (26)$$

$$\sum_{e'' \in \text{In}(v)} \beta_{e,e''}^2(t) = 1. \quad (27)$$

In [36], we established the relation between the satisfaction of RIP and the tail probability

$$\begin{aligned} \mathbf{p}_{\text{tail}}(\Psi_{\text{tot}}(t), \varepsilon) &= \max_{\underline{x}'} \mathbf{P} \left(\left| \|\Psi_{\text{tot}}(t) \underline{x}'\|_{\ell_2}^2 - 1 \right| > \varepsilon \right), \\ &\text{subject to: } \|\underline{x}'\|_{\ell_2} = 1 \end{aligned} \quad (28)$$

by proving the following theorem.

Theorem 4.3: (Theorem 4.1 in [36]) Consider $\Psi_{\text{tot}}(t)$ with the tail probability, as defined in (28), and an orthonormal transform matrix ϕ . Then, $\Theta_{\text{tot}}(t) = \Psi_{\text{tot}}(t) \cdot \phi$ satisfies RIP of order k and constant δ_k , with a probability exceeding,

$$1 - \binom{n}{k} \left(\frac{42}{\delta_k} \right)^k \mathbf{p}_{\text{tail}}(\Psi_{\text{tot}}(t), \varepsilon = \frac{\delta_k}{\sqrt{2}}). \quad \blacksquare \quad (29)$$

By using theorem 4.3, we can analyze the behavior of $\Psi_{\text{tot}}(t)$, resulting from the proposed local network coding coefficients in theorem 4.2. Specifically, we try to see if we can obtain the same tail probability as a Gaussian ensemble, with the same order of measurements. Unfortunately, the complicated relation of local network coding coefficients and network parameters with the resulting $\Psi_{\text{tot}}(t)$ (see Eqs. 8, 9,

16, 21) makes it difficult to derive a simple mathematical form for the tail probability, $\mathbf{p}_{\text{tail}}(\Psi_{\text{tot}}(t), \varepsilon)$, and have a nice mathematical conclusion about the required number of measurements.

In Fig. 3, we present the numerical values of tail probabilities (defined in Eq. 28) for the resulting $\Psi_{\text{tot}}(t)$, $\mathbf{p}_{\text{tail}}(\Psi_{\text{tot}}(t), \varepsilon)$, using the proposed local network coding coefficients in theorem 4.2. These tail probabilities are compared with those of i.i.d Gaussian matrices, \mathbf{G} , versus the number of measurements, m , in each case.⁴

Remark 4.4: Our numerical evaluations in Fig. 3 show that for the same value of tail probability, there is a QNC resulting measurement matrix, $\Psi_{\text{tot}}(t)$, and an i.i.d Gaussian matrix, \mathbf{G} , which have the same order of measurements, m . Furthermore, using theorem 4.3, we can say the the resulting $\Psi_{\text{tot}}(t)$ has a similar behavior as Gaussian matrices, in terms of RIP satisfaction.

In the following section, we use the aforementioned conclusion for the resulting $\Psi_{\text{tot}}(t)$ to derive a performance bound for QNC scenario.

V. DECODING USING SPARSE RECOVERY

Sparse recovery for exactly sparse data can be done by using linear programming [37], where NP-hard ℓ_0 minimization is replaced by ℓ_1 minimization. Fortunately, this alteration of cost function does not affect the optimality in recovery of exactly sparse vectors from noiseless measurements [32], [37]. However, when dealing with noisy measurements, ℓ_1 -min recovery does not necessarily offer an optimal solution. There is still a lot of work being done to develop practical and near-optimal recovery algorithms for noisy cases. In the following, we discuss ℓ_1 -min recovery for QNC scenario and establish theoretical bounds on its recovery error.

Motivated by the work in [20], [33], the compressed sensing based decoder for QNC scenario solves the following convex optimization:

$$\begin{aligned} \hat{\underline{X}}(t) &= \phi \cdot \arg \min_{\underline{S}'} \|\underline{S}'\|_{\ell_1}, \\ \text{subject to: } &\|\underline{Z}_{\text{tot}}(t) - \Psi_{\text{tot}}(t) \phi \underline{S}'\|_{\ell_2} \leq \epsilon_{\text{rec}}(t) \end{aligned} \quad (30)$$

which can be solved by using linear programming [37]. In the following, we present our results on the recovery error using ℓ_1 -min decoding of Eq. 30.

Theorem 5.1: Consider a QNC scenario where, for all $v \in \mathcal{V}$, the network coding coefficients satisfy

⁴Detailed version of our calculations for the tail probability of $\Psi_{\text{tot}}(t)$ can be found in [36].

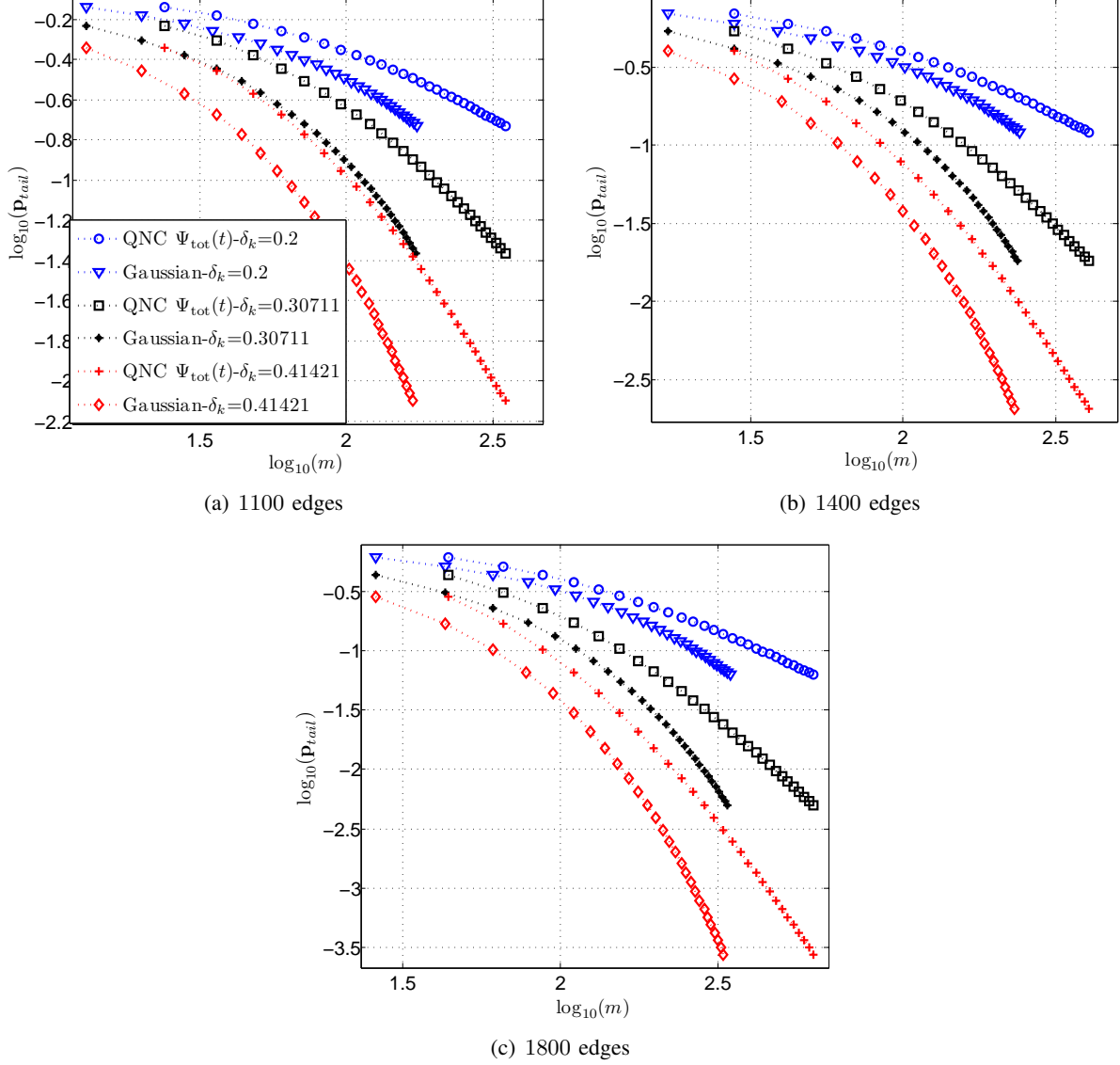


Fig. 3. Logarithmic tail probability versus logarithmic ratio of minimum required number of measurements in our QNC scenario and i.i.d. Gaussian measurement matrices, for $n = 100$ nodes, different RIP constants, and different number of edges .

the following normalization condition:

$$\sum_{e' \in In(v)} |\beta_{e,e'}(t)| + |\alpha_{e,v}(t)| \leq 1, \quad \forall e \in Out(v), \quad \forall t. \quad (31)$$

In such scenario, we assume that the resulting $\Theta_{\text{tot}}(t) = \Psi_{\text{tot}}(t)\phi$ satisfies RIP of order $2k$ with constant δ_{2k} , where $\delta_{2k} < \sqrt{2} - 1$. Moreover, the messages, X_v 's, are supposed to be bounded between $-q_{\max}$

$$\epsilon_{rec}^2(t) = \frac{1}{4} \sum_{t'=2}^t \left(\sum_{t''=1}^{t'-1} \underline{\Delta}_Q^T \left| \prod_{t'}^{t'''=t''+2} F(t''') \right| \cdot B^T B \cdot \sum_{t''=1}^{t'-1} \left| \prod_{t'}^{t'''=t''+2} F(t''') \right| \underline{\Delta}_Q \right) \quad (33)$$

and $+q_{max}$, and the edge quantizers, $\mathbf{Q}_e(\cdot)$'s, are uniform with the step size

$$\Delta_e = \frac{2q_{max}}{[2^{LC_e}]}. \quad (32)$$

Now, for the ℓ_1 -min decoding of (30) where $\epsilon_{rec}^2(t)$ is as defined in Eq. 33, and $\underline{\Delta}_Q = [\Delta_e : e \in \mathcal{E}]$, we have:

$$\left\| \hat{\underline{X}}(t) - \underline{X} \right\|_{\ell_2} \leq c_1 \epsilon_{rec} + \frac{c_2}{\sqrt{2}} \epsilon_k. \quad (34)$$

In the inequality of (34), c_1 and c_2 , are constants, defined as follows:

$$c_1 = 4 \frac{\sqrt{1 + \delta_{2k}}}{1 - (1 + \sqrt{2})\delta_{2k}}, \quad (35)$$

$$c_2 = 2 \frac{1 - (1 - \sqrt{2})\delta_{2k}}{1 - (1 + \sqrt{2})\delta_{2k}}. \quad (36)$$

Proof: Since the network is lossless and network coding coefficients satisfy the condition of Eq. 31, and $|X_v| \leq q_{max}$, $\forall v$, the only associated measurement noise is resulting from the quantization noise at the edges. Moreover, for each $e \in \mathcal{E}$, we have:

$$-\frac{\Delta_e}{2} \leq N_e(t) \leq +\frac{\Delta_e}{2}, \quad (37)$$

since the quantizers are uniform. Equivalently, the absolute value vector of $\underline{N}(t)$, represented by $|\underline{N}(t)|$, is such that:

$$|\underline{N}(t)| \leq \frac{1}{2} \underline{\Delta}_Q. \quad (38)$$

Since B is a one-to-one mapping matrix with positive entries, the effective noise vector, $\underline{N}_{\text{eff}}(t)$, can be upper bounded as follows:

$$|\underline{N}_{\text{eff}}(t)| = \left| B \sum_{t'=1}^{t-1} \prod_t^{t''=t'+2} F(t'') \underline{N}(t'+1) \right| \quad (39)$$

$$= B \cdot \left| \sum_{t'=1}^{t-1} \prod_t^{t''=t'+2} F(t'') \underline{N}(t'+1) \right|. \quad (40)$$

Therefore, using (38), we have:

$$\left| \underline{N}_{\text{eff}}(t) \right| \leq B \cdot \sum_{t'=1}^{t-1} \left| \prod_t^{t''=t'+2} F(t'') \underline{N}(t'+1) \right| \quad (41)$$

$$\leq B \cdot \sum_{t'=1}^{t-1} \left| \prod_t^{t''=t'+2} F(t'') \right| \cdot \left| \underline{N}(t'+1) \right| \quad (42)$$

$$\leq B \cdot \sum_{t'=1}^{t-1} \left| \prod_t^{t''=t'+2} F(t'') \right| \cdot \frac{1}{2} \Delta_Q. \quad (43)$$

Finally, by using inequality of (43),

$$\left\| \underline{N}_{\text{eff,tot}}(t) \right\|_{\ell_2}^2 = \sum_{t'=2}^t \left\| \underline{N}_{\text{eff}}(t') \right\|_{\ell_2}^2, \quad (44)$$

and,

$$\left\| \underline{N}_{\text{eff}}(t') \right\|_{\ell_2}^2 = \left| \underline{N}_{\text{eff}}(t) \right|^T \cdot \left| \underline{N}_{\text{eff}}(t) \right|, \quad (45)$$

we can show that:

$$\left\| \underline{N}_{\text{eff,tot}}(t) \right\|_{\ell_2}^2 \leq \epsilon_{\text{rec}}^2(t), \quad (46)$$

such that $\epsilon_{\text{rec}}(t)$ is as defined in Eq. 33. Now, by applying theorem 4.2 in [21], and using the definition of ϵ_k in (4), we can finish the proof of our theorem. ■

According to the preceding theorem, the upper bound, $c_1 \epsilon_{\text{rec}}$, is decreased when the quantization steps, Δ_e 's, are decreased, too. And since $\Delta_e = 2q_{\text{max}} / \lfloor 2^{L C_e} \rfloor$, a smaller upper bound on the ℓ_2 -norm of recovery error is forced by increasing the block length, L . Although this can be done practically, it will simultaneously increase the point to point transmission delays in the network, which may not be desirable. Introducing a trade-off on the choice of block length, one has to find its appropriate value for a specific quality of service (*i.e.* recovery error).

Based on theorem 5.1, if the resulting $\Psi_{\text{tot}}(t)$ satisfies RIP of appropriate order with a high probability, then the robust recovery can be guaranteed. On the other hand, using remarks 4.1,4.4, we can say that the resulting $\Psi_{\text{tot}}(t)$ satisfies RIP with a high probability, while the number of measurements, m , has a smaller order than the number of messages, n . Therefore, putting all these numerical and theoretical results together, it is true to say that QNC can result in bounded error recovery (34) with a smaller order of measurements (received packets at the decoder) than that of messages. This saving in the required number of received packets can be interpreted as an *embedded distributed compression*, achieved by quantized network coding at the nodes.

TABLE I
THE PARAMETERS OF MESSAGES AND THE NETWORKS, USED IN OUR SIMULATIONS.

Parameter	Value(s)
No of nodes, n	100
No of edges, $ \mathcal{E} $	1100, 1400, 1800
Block length, L	$1, \dots, 40$
Sparsity factor, k/n	0.05, 0.15, 0.25
Near-sparsity factor, $\epsilon_k/\ \underline{s}_k\ _{\ell_1}$	0, 0.002, 0.02, 0.2
Range of messages	$[-10, +10]$
Average $\ \underline{x}\ _{\ell_2}$	$41.7 \sim 32.4$ [dB]

VI. SIMULATION RESULTS

In this section, we evaluate the performance of quantized network coding, by using different numerical simulations. We are interested to find out the compression achievements, resulting from QNC, by obtaining the delay-distortion curves in different scenarios.

Although we were able to derive mathematical performance measures for the QNC scenario, they are not comprehensive and do not offer any guarantee on the statistical performance measures; *e.g.* mean squared error. However, deriving such statistical performance bounds requires a lot more of theoretical work on the sparse recovery, and meanwhile we can only rely on the numerical evaluations.

We initiate our numerical evaluations, by comparing the delay-quality performance of QNC and conventional routing based packet forwarding for lossy transmission of a set of correlated sources (messages). To set up the simulations, we randomly generate random deployments of directed networks with uniformly distributed edges (making sure there is not any pair of nodes with two assigned edges). The edges can maintain a lossless communication of one bit per use, meaning $C_e = 1$, for all $e \in \mathcal{E}$. One of the nodes is randomly picked to be the gateway node, v_0 , in which the messages are decoded. To generate a realization of messages, \underline{x} , we first generate a k -sparse random vector, \underline{s}_k , whose components are uniformly distributed between $-\frac{1}{2}$ and $+\frac{1}{2}$. A near-sparse vector, \underline{s} , is obtained by adding a zero mean uniform noise, such that $\|\underline{s} - \underline{s}_k\|_{\ell_2}$ is bounded by ϵ_k . This is followed by generation of an orthonormal random matrix, ϕ , and calculating random messages; $\underline{x} = \phi \cdot \underline{s}$, and normalizing the range of x_v 's, between $-q_{max}$ and $+q_{max}$, where $q_{max} = 10$. Different values of sparsity factor, $\frac{k}{n}$, and ϵ_k are used in our simulations. A summary of the simulation parameters is presented in Table I.

For each generated random network deployment, we perform QNC with ℓ_1 -min decoding. Local net-

work coding coefficients, $\alpha_{e,v}(t)$'s and $\beta_{e,e'}(t)$'s, are generated according to the conditions of theorem 4.2. The freedom degrees are limited by picking $\beta_{e,e'}(t)$'s such that they are locally orthonormal; The resulting coefficients are then normalized to satisfy the normalization condition of Eq. 31 and prevent overflow in the linear combination of QNC. Edge quantizers, $\mathbf{Q}_e(\cdot)$'s, are uniform with a range of $[-q_{max}, +q_{max}]$ and 2^L intervals (since $C_e = 1, \forall e$). This completes all the required parameters and vectors to simulate quantized network coding and obtain the received packets at the decoder node, $\underline{z}(t)$'s.⁵ Random $\alpha_{e,v}(2)$'s can be generated in a pseudo-random way and therefore only the generator seed needs to be transmitted to the decoder as a header.

At the decoder, the received measurements up to t , $\underline{z}_{\text{tot}}(t)$, are used to recover the original messages. Specifically, for a realization of messages, \underline{x} , we define $\hat{\underline{x}}_{\text{QNC}}(t)$, to be the recovered messages, using ℓ_1 -min decoding, according to (30). Moreover, the convex optimization, involved in (30) is solved by using the open source implementation of disciplined convex programming [38], [39].

For each deployment, we also simulated a routing based packet forwarding and compare it with the results for QNC. To find the routes from each node to the gateway node, we find the shortest path from each node to the gateway node, using the Dijkstra algorithm [40]. Further, the real valued messages, x_v 's, are quantized at their corresponding source nodes, by using similar uniform quantizers, as used in QNC transmission. It is aimed to deliver all x_v 's to the decoder node and keep the track of delivered messages over time, t , in the recovered vector of messages, $\hat{\underline{x}}_{\text{PF}}(t)$. Moreover, if a message, x_v , is not delivered by time index t , zero is used as its recovered value:

$$\{\hat{\underline{x}}_{\text{PF}}(t)\}_v = 0. \quad (47)$$

The ℓ_2 norm of recovery error, $\|\underline{x} - \hat{\underline{x}}(t)\|_{\ell_2}$, is used as the quality measure in our numerical comparisons. The payback measure in our comparisons is the delivery delay, corresponding to achieve a minimum quality of service. Explicitly, delivery delay for a transmission which has terminated at t is equal to $(t-1)L$ in both cases of QNC and packet forwarding.⁶ In QNC scenario, for each value of $\frac{k}{n}$, and ϵ_k , we calculate the average of $\|\underline{x} - \hat{\underline{x}}_{\text{QNC}}(t)\|_{\ell_2}$'s over different realizations of network deployments. Since the sparsity of messages does not affect the performance of packet forwarding, we only need to present its results for different network parameters (*i.e.* number of edges).

For a fixed block length, $L = 20$, the average of ℓ_2 norm of recovery error versus the average delivery

⁵Lower case notations are used for realization of random variables.

⁶In the case of packet forwarding, we do not consider the learning period, required to find the optimal routes.

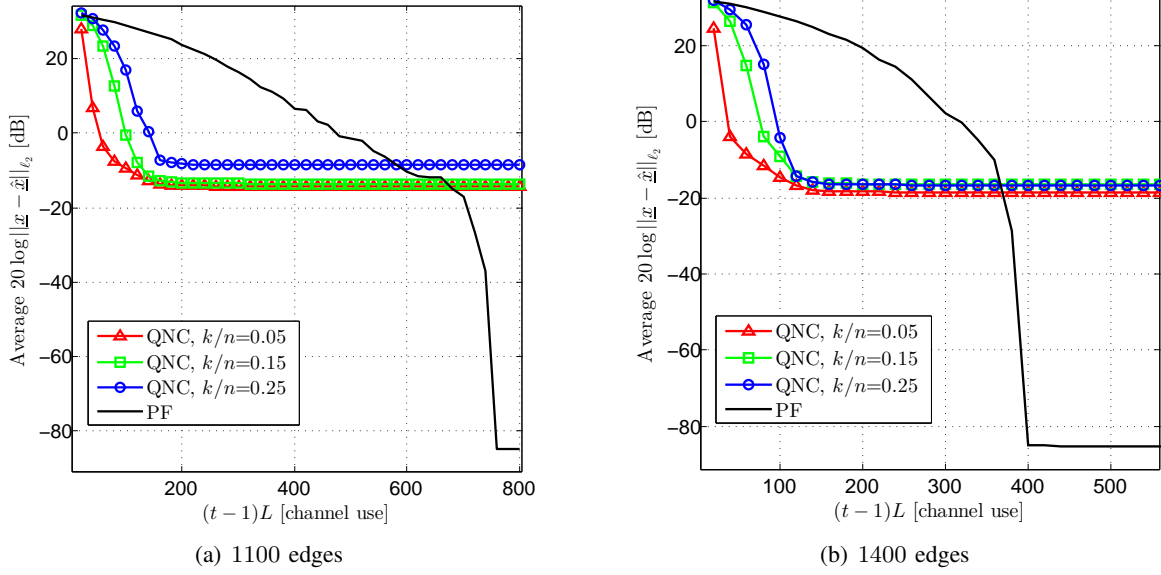


Fig. 4. Average ℓ_2 -norm of recovery error versus delivery delay for QNC and PF scenarios, when $L = 20$, $\epsilon_k = 0$ and $|\mathcal{E}| = 1100, 1400$.

delay is depicted in Fig. 4. In Figs. 4(b),4(a), the horizontal axis represents the product $(t-1)L$, which is the delivery delay, corresponding to $L = 20$, for different values of $t \geq 1$. The vertical axis is the logarithmic average ℓ_2 norm of recovery error, $20 \log_{10} \|\underline{x} - \hat{\underline{x}}\|_{\ell_2}$, for QNC and Packet Forwarding (PF) scenarios. Simply, Figs. 4(b),4(a) can be considered as rate-distortion curves in a lossy source coding scenario.

As it is shown in Figs. 4(b),4(a), when using the same block length, QNC achieves significant improvement, compared to PF, for low values of delivery delay. These low delays correspond to the initial t 's in the transmission, at which a small number of packets are received at the decoder, as expected. After enough packets are received at the decoder, QNC achieves its best performance (at around -20 [dB]), as a result of associated quantization noises. The best performance for packet forwarding happens after a longer period of time than for QNC. On the other hand, the best performance of PF (around -80 [dB]) is higher than that of QNC, which can be explained by noise propagation in the network during QNC. However, as it is shown in the following, QNC outperforms PF in a wide range of delay values, when the appropriate block length is adopted in each case.

After simulating QNC and PF scenarios for different block lengths and calculating the corresponding delay and recovery error norms, we find the best values of block length for each specific average ℓ_2 norm of recovery error (as a measure of quality of service). The resulting L -optimized curve for each QNC and

PF scenario is depicted in Fig. 5. In Figs. 5(a)-5(f), QNC performance is compared with that of PF, for different number of edges, different sparsity factors, and near-sparsity parameters. Generally speaking, these figures show a promising improvement over conventional packet forwarding, when QNC is adopted for transmission of near-sparse messages. The achieved improvement is increased as the sparsity factor, $\frac{k}{n}$, is decreased, meaning a higher level of correlation between the messages.

As a drawback, QNC seems to fail when the sparsity model is not good for describing the correlation model. Specifically, if the near-sparsity parameter, ϵ_k , is too high, then the resulting performance of QNC can not even achieve that of PF, for a wide range of delivery delays (see Fig. 5(b) for instance). In Figs. 6(a),6(b),6(c), the effect of ϵ_k on the resulting QNC performance is illustrated. As it is shown, as long as ϵ_k is small (relative to the ℓ_1 norm of message), there is not any difference in the QNC performance. But, if it is so high that the sparsity model does not characterize the messages fairly, then QNC fails to work properly (since ℓ_1 min decoding criteria is not a good cost function anymore).

In the routing based packet forwarding scenarios, the intermediate (sensor) nodes have to go through route training and storage of packets. As one of the main advantages of network coding, in QNC scenario, the intermediate nodes should only carry simple linear combination and quantization, being liberated in terms of computational complexity. On the other side, at the decoder sides, QNC requires an ℓ_1 -min decoder which is potentially more complex than the receiver required for packet forwarding. However, this may not be an issue in practical cases, as the gateway node is usually capable of handling high computational operations.

VII. CONCLUSIONS AND FUTURE WORKS

Joint source network coding of correlated sources was studied with a sparse recovery perspective. In order to achieve non-adaptive encoding, we proposed quantized network coding, which incorporates real field network coding and quantization to take advantage of decoding using linear programming. Thanks to the work in the literature of compressed sensing, we discussed theoretical guarantees to ensure efficient encoding and robust decoding of messages. Moreover, we were able to make conclusive statements about the robust recovery of messages, when fewer number of received packets than the number of sources (messages) were available at the decoder. Finally, our computer simulations verified the reduction in the average delivery delay, by using quantized network coding.

Although the proposed sparse recovery algorithm is working well for correlated messages with near-sparse characterization, it does not offer optimal recovery for other cases of correlated sources. Currently, we are studying the feasibility of near-optimal decoding, when other forms of prior information are

available about the source. Specifically, we have suggested the use of belief propagation based decoding [41], [42] in a Bayesian scenario. However, more theoretical work is needed to derive mathematical guarantees for robust recovery. Studying the general case of lossy networks with interference between the links is also one of the urging needs for our work.

REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.
- [2] C. Chong and S. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.
- [3] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, July 2000.
- [4] J. Al-Karaki and A. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, 2004.
- [5] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.
- [6] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 80–94, 2004.
- [7] T. S. Han, "Slepian-wolf-cover theorem for networks of channels," *Information and Control*, vol. 47, no. 1, pp. 67–83, 1980.
- [8] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *IEEE International Symposium on Information Theory*, p. 442, June–4 July 2003.
- [9] C. Fragouli, "Network coding for sensor networks," *Handbook on Array Processing and Sensor Networks*, pp. 645–667, 2009.
- [10] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE Transactions on Networking*, vol. 11, no. 5, pp. 782–795, 2003.
- [11] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.
- [12] S. Lim, Y. Kim, A. El Gamal, and S. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132–3152, 2011.
- [13] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 789–804, 2006.
- [14] T. Ho, M. Médard, M. Effros, R. Koetter, and D. Karger, "Network coding for correlated sources," in *Proceedings of Conference on Information Sciences and Systems*, 2004.
- [15] A. Ramamoorthy, K. Jain, P. A. Chou, and M. Effros, "Separating distributed source coding from network coding," *IEEE Transactions on Networking*, vol. 14, pp. 2785–2795, June 2006.
- [16] Y. Wu, V. Stankovic, Z. Xiong, and S. Kung, "On practical design for joint distributed source and network coding," *IEEE Transactions on Information Theory*, vol. 55, no. 4, pp. 1709–1720, 2009.

- [17] G. Maierbacher, J. Barros, and M. Médard, “Practical source-network decoding,” in *6th International Symposium on Wireless Communication Systems*, pp. 283–287, IEEE, 2009.
- [18] S. Cruz, G. Maierbacher, and J. Barros, “Joint source-network coding for large-scale sensor networks,” in *IEEE International Symposium on Information Theory Proceedings*, pp. 420–424, IEEE, 2011.
- [19] F. Kschischang, B. Frey, and H. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [20] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, April 2006.
- [21] R. Baraniuk, M. Davenport, M. Duarte, and C. Hegde, *An Introduction to Compressive Sensing*. Addison-Wesley, 2011.
- [22] J. Haupt, W. Bajwa, M. Rabbat, and R. Nowak, “Compressed sensing for networked data,” *IEEE Signal Processing Magazine*, vol. 25, pp. 92–101, march 2008.
- [23] N. Nguyen, D. Jones, and S. Krishnamurthy, “Netcompress: Coupling network coding and compressed sensing for efficient data communication in wireless sensor networks,” in *2010 IEEE Workshop on Signal Processing Systems*, pp. 356–361, oct. 2010.
- [24] C. Luo, F. Wu, J. Sun, and C. W. Chen, “Compressive data gathering for large-scale wireless sensor networks,” in *Proceedings of the 15th annual international conference on Mobile computing and networking*, MobiCom ’09, (New York, NY, USA), pp. 145–156, ACM, 2009.
- [25] S. Feizi, M. Médard, and M. Effros, “Compressive sensing over networks,” in *48th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1129–1136, IEEE, 2010.
- [26] W. Xu, E. Mallada, and A. Tang, “Compressive sensing over graphs,” in *IEEE International Conference on Computer Communications (INFOCOM)*, pp. 2087–2095, IEEE, 2011.
- [27] M. Wang, W. Xu, E. Mallada, and A. Tang, “Sparse recovery with graph constraints: Fundamental limits and measurement construction,” in *IEEE International Conference on Computer Communications (INFOCOM)*, pp. 1871–1879, IEEE, 2012.
- [28] S. Feizi and M. Medard, “A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing,” in *49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1048–1054, IEEE, 2011.
- [29] F. Bassi, L. Chao, L. Iwaza, M. Kieffer, *et al.*, “Compressive linear network coding for efficient data collection in wireless sensor networks,” in *Proceedings of the 2012 European Signal Processing Conference*, pp. 1–5, 2012.
- [30] M. Nabaee and F. Labeau, “Quantized network coding for sparse messages,” *arXiv preprint arXiv:1201.6271*, 2012.
- [31] T. Kailath, *Linear systems*, vol. 1. Prentice-Hall Englewood Cliffs, NJ, 1980.
- [32] E. Candes and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse problems*, vol. 23, no. 3, p. 969, 2007.
- [33] E. J. Cands, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathématique*, vol. 346, no. 9-10, pp. 589–592, 2008.
- [34] R. Baraniuk, “Compressive sensing,” *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.
- [35] R. Baraniuk, M. Davenport, R. Devore, and M. Wakin, “A simple proof of the restricted isometry property for random matrices,” *Constr. Approx.*, vol. 2008, 2007.
- [36] M. Nabaee and F. Labeau, “Restricted isometry property in quantized network coding of sparse messages,” *arXiv preprint arXiv:1203.1892*, 2012.
- [37] E. Candes and T. Tao, “Decoding by linear programming,” *IEEE Transactions on Information Theory*, vol. 51, pp. 4203–4215, December 2005.

- [38] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 1.21.” <http://cvxr.com/cvx>, 2011.
- [39] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control* (V. Blondel, S. Boyd, and H. Kimura, eds.), Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited, 2008.
- [40] E. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.
- [41] M. Nabae and F. Labeau, “Bayesian quantized network coding via belief propagation,” *arXiv preprint arXiv:1209.1679*, 2012.
- [42] M. Nabae and F. Labeau, “One-step quantized network coding for near sparse messages,” *arXiv preprint arXiv:1210.7399*, 2012.

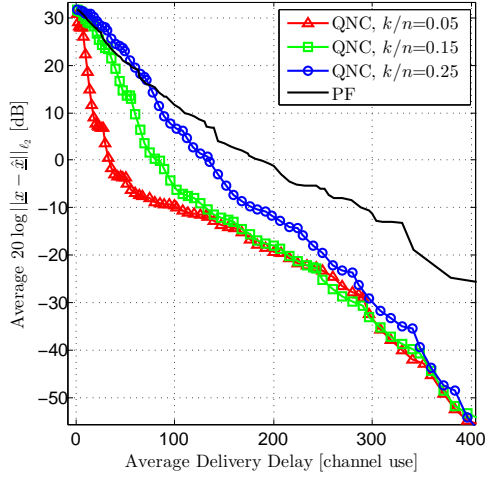
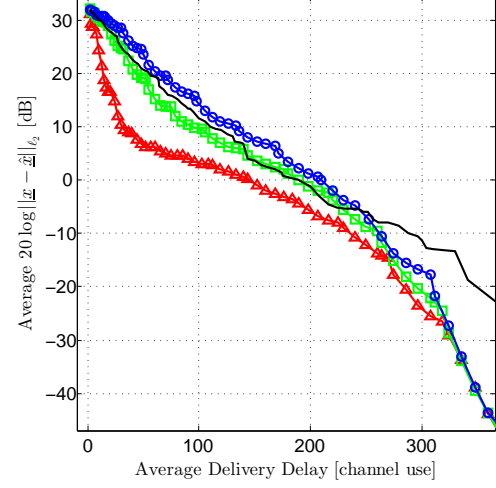
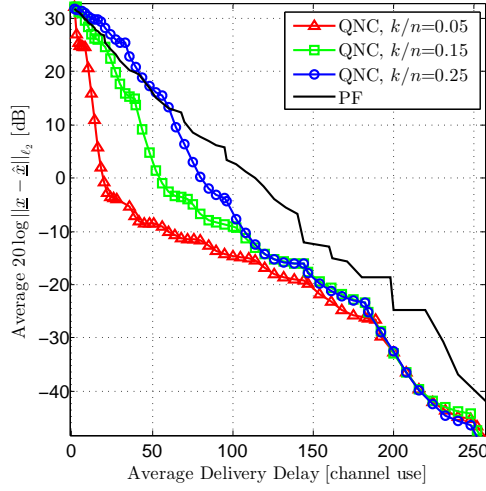
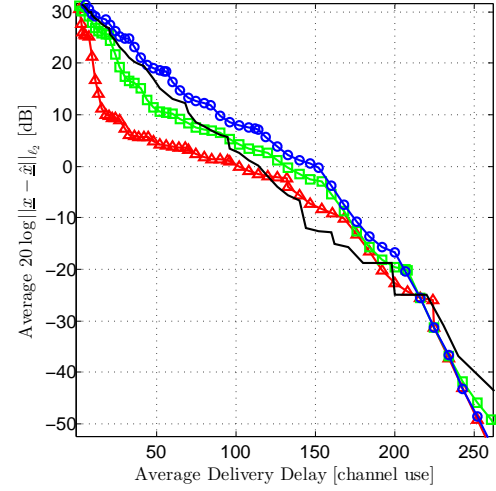
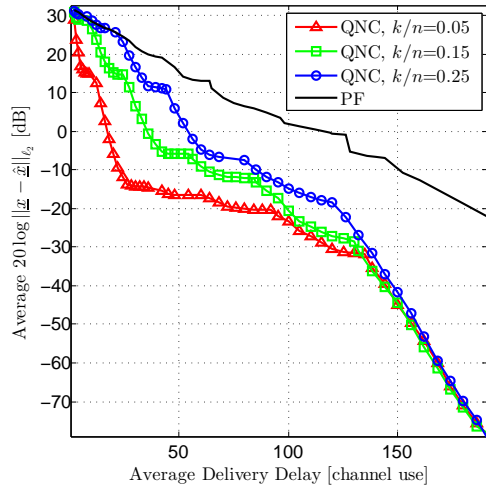
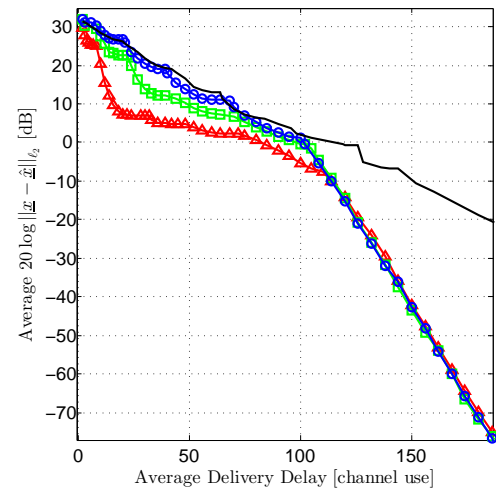
(a) 1100 edges, $\epsilon_k = 0$ (b) 1100 edges, $\epsilon_k = 0.2$ (c) 1400 edges, $\epsilon_k = 0$ (d) 1400 edges, $\epsilon_k = 0.2$ (e) 1800 edges, $\epsilon_k = 0$ (f) 1800 edges, $\epsilon_k = 0.2$

Fig. 5. Average ℓ_2 -norm of recovery error versus average delivery delay for QNC and PF, and different values of ϵ_k .

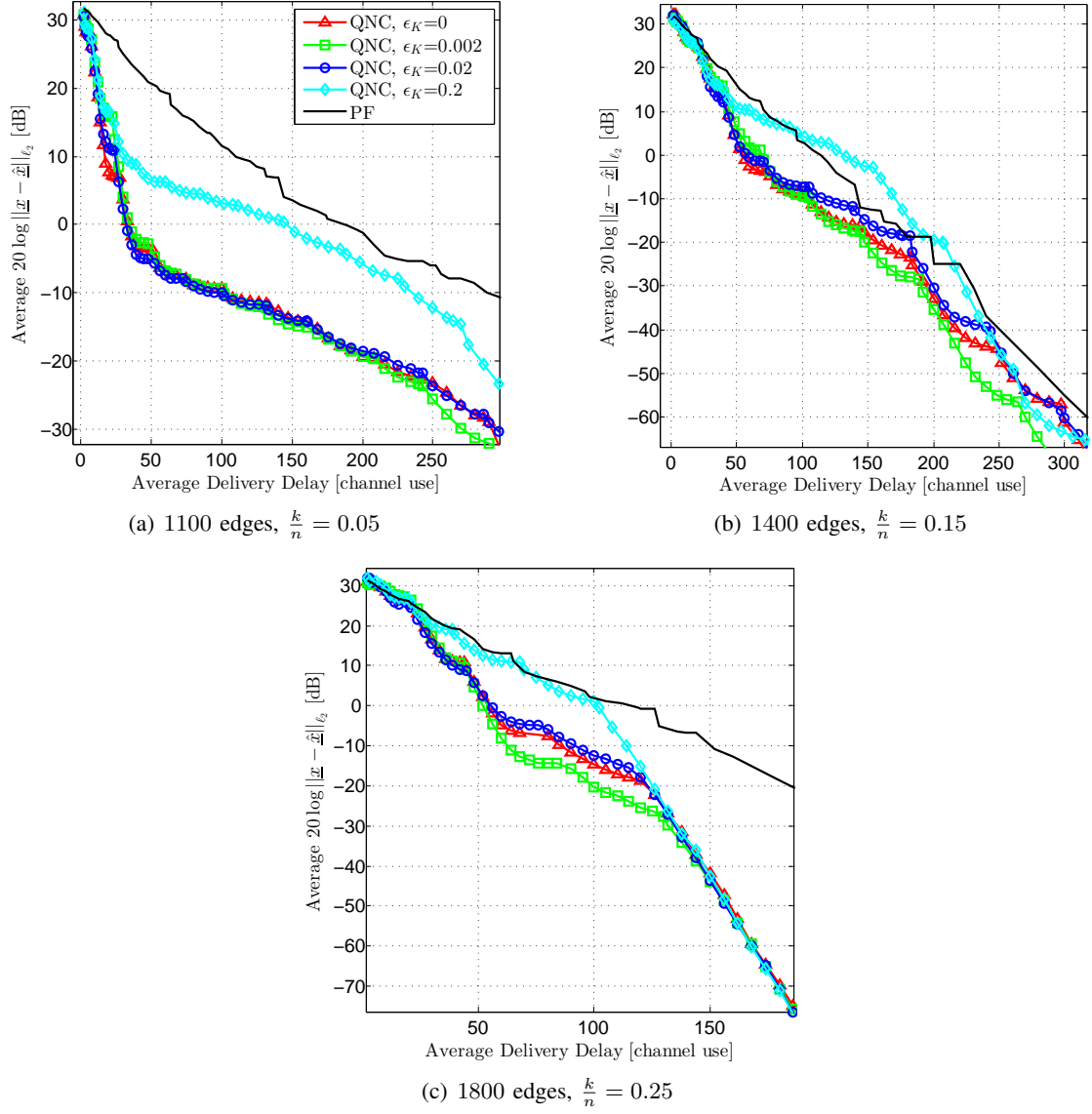


Fig. 6. Comparison of QNC versus PF performance for different near-sparsity parameters, ϵ_k .